JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 57 (1995) 403–412

Rigorous perturbation bounds for eigenvalues and eigenvectors of a matrix

A.S. Deif

Department of Engineering Mathematics, Faculty of Engineering, Cairo University, Giza, Egypt

Received 10 February 1993; revised 6 July 1993

Abstract

Tight perturbation bounds are given for the shifts in the eigenvalues and eigenvectors of a matrix. The case of simple as well as multiple eigenvalues is treated with the results tested numerically.

Keywords: Perturbation bounds; Error estimates; Accuracy of matrix eigensystem; Matrix polynomial

1. Introduction

Perturbation theory for matrices attracted great interest due to its applicability in many fields. The theory itself is quite understood [2, 8, 14] and yet there is room for trying to furnish rigorous bounds as a method for crudely approximating the error in the computed eigenvalues and eigenvectors. If A has the eigenpair (λ, x), where λ stands for the eigenvalue and x for the eigenvector, then it is of interest to obtain bounds for the shifts $|\tilde{\lambda} - \lambda|$ and $|\tilde{x} - x|$, where $(\tilde{\lambda}, \tilde{x})$ is the eigenpair of $A + \delta A$, in terms of the perturbation δA. Apart from applications in engineering like sensitivity analysis of λ and x to changes in the matrix elements, such bounds are useful in numerical analysis if $\delta A = \varepsilon A_1$ with ε corresponding to the machine eps.

Classical perturbation theory suggests that $(\tilde{\lambda}, \tilde{x})$ be expanded in a power series of ε, namely

$$\begin{aligned}\tilde{\lambda}(\varepsilon) &= \lambda + \varepsilon\lambda^{(1)} + \varepsilon^2\lambda^{(2)} + \cdots\\ \tilde{x}(\varepsilon) &= x + \varepsilon x_{(1)} + \varepsilon^2 x_{(2)} + \cdots\end{aligned} \tag{1}$$

for λ simple (Puiseux series appear if λ is multiple [8]). And by equating equal powers of ε in the operator equation

$$(A + \varepsilon A_1)\tilde{x}(\varepsilon) = \tilde{\lambda}(\varepsilon)\tilde{x}(\varepsilon) \tag{2}$$

one can obtain recursively $\lambda^{(1)}$, $x_{(1)}$, $\lambda^{(2)}$, etc. [6, 10, 11]. It is to be noted that, practically speaking, $\tilde{\lambda}(\varepsilon)$ and $\tilde{x}(\varepsilon)$ are hardly required; or else, with the advent of fast computers one can perform several

SSDI 0377-0427(93)E0208-4

runs for different values of ε and determine a rough estimate of them. Although such trends frequently used by many workers can be very misleading as Kulisch and Miranker reported [9], it still gives an idea about the amount of ill-conditioning in the eigensystem.

On the contrary, perturbation theory is mainly used to calculate $\lambda^{(1)}$ and $x_{(1)}$. This has applications in:

(1) sensitivity analysis, like when it is needed to calculate $\partial\lambda_i/\partial a_{kj}$ [4, 12, 15];

(2) numerical analysis, $\lambda^{(1)}$ is an indication of the matrix conditioning [13, 16].

Perturbation theory is also not void of disadvantages. Of the latter, we mention:

(1) The expansion (1) must first be proved to exist; for sometimes Puiseux expansions occur, in which case, $|\tilde{\lambda} - \lambda| \approx \mathrm{O}(\varepsilon^{1/m})$, where m is the multiplicity of λ.

(2) Expansion (1) must be shown to be convergent for sufficiently small $|\varepsilon|$, otherwise $\lambda^{(1)}$ cannot be taken as roughly representing the shift in λ. Convergence criteria are usually difficult to find, let alone applied (Gerschgorin's disk theorem, etc.). Therefore, unless ε is very small (e.g. $\varepsilon \approx \mathrm{eps(machine)}$), $\lambda^{(1)}$ cannot bound the shift in λ.

(3) The case ε is only known in an interval, i.e. only a lower and upper bound for δA are at hand.

In this work, we shall try to provide strict bounds for $|\tilde{\lambda}_i - \lambda_i|$ and $|\tilde{x}^i - x^i|$ which are valid for δA big. The only restriction is that the separation between the eigenvalues is expected not to exceed their shifts.

Among the first attempts, however, to obtain global bounds for eigenvalue shifts is the Bauer–Fike one [1]. They showed that the eigenvalues of $A + \delta A$ lie in the union of the disks

$$M_i = \{z\colon |z - \lambda_i| \leqslant \|T^{-1}\|\,\|T\|\,\|\delta A\|\}, \quad i = 1, \ldots, n, \tag{3}$$

where $\|\cdot\|$ is some norm and T is the modal matrix containing the eigenvectors of A and which brings it to a diagonal form under a similarity transformation; namely $T^{-1}AT = \Lambda$. The bound in (3) follows easily upon noticing that since

$$\tilde{\lambda}I - A - \delta A = (\tilde{\lambda}I - A)[I - (\tilde{\lambda}I - A)^{-1}\delta A] \tag{4}$$

is singular, then $\|(\tilde{\lambda}I - A)^{-1}\delta A\| > 1$. And by writing $(\tilde{\lambda}I - A)^{-1} = T(\tilde{\lambda}I - \Lambda)^{-1}T^{-1}$, one has that

$$1 \leqslant \|T(\tilde{\lambda}I - \Lambda)^{-1}T^{-1}\delta A\| \leqslant \|T\|\,(1/\min|\tilde{\lambda} - \lambda|)\,\|T^{-1}\|\,\|\delta A\| \tag{5}$$

from which (3) follows. However, unless A is normal ($\|T\| = 1$ under $\|\cdot\|_2$), (3) is known to yield pessimistic results (see example in Section 5).

2. New bound

From (4), having that $\tilde{\lambda}I - A - \delta A$ is singular, whereas $\det(\tilde{\lambda}I - A) \neq 0$,

$$\begin{aligned} 0 &= \det(I - (\tilde{\lambda}I - A)^{-1}\delta A) = \det(I - T(\tilde{\lambda}I - \Lambda)^{-1}T^{-1}\delta A) \\ &= \det(I - (\tilde{\lambda}I - \Lambda)^{-1}T^{-1}\,\delta A\,T) \end{aligned} \tag{6}$$

implying for some vector $z \neq 0$, that

$$|z| \leqslant |(\tilde{\lambda}I - \Lambda)^{-1}||T^{-1}||\delta A||T||z| \tag{7}$$

where $|\cdot|$ denote absolute values taken componentwise. Thus

$$\|z\| \leqslant (1/\min|\tilde{\lambda} - \lambda|)\||T^{-1}||\delta A||T|\|\,\|z\| \tag{8}$$

or that

$$\min|\tilde{\lambda} - \lambda| \leqslant \||T^{-1}||\delta A||T|\| \tag{9}$$

valid for any δA. In floating-point computation $|\delta A| \leqslant \text{eps}|A|$ and (9) reduces to the practical bound

$$|\tilde{\lambda} - \lambda| \leqslant \text{eps}\,\||T^{-1}||A||T|\|. \tag{10}$$

Apart from being tighter than the Bauer–Fike bound especially for nonnormal matrices, the bound in (10) is scale-invariant under the transformation $A = DBD^{-1}$ where D is diagonal. Thus (3) is loose for a matrix similar to a symmetric matrix under the foregoing transformation or for a matrix A similar to a matrix B having a well-conditioned eigenproblem.

3. Bounds for A defective

The inequality in (7) is valid, except that we shall take for Λ the Jordan form of A called J, and T will contain the generalized eigenvectors. Let A have an eigenvalue λ with multiplicity m and index r (by index we mean the size of the largest Jordan block associated with λ). Then under the ℓ_∞ or ℓ_1-norm

$$\|(\tilde{\lambda}I - J)^{-1}\| = \sum_{i=1}^{r} 1/|\tilde{\lambda} - \lambda|^i \leqslant r\max(1/|\tilde{\lambda} - \lambda|, 1/|\tilde{\lambda} - \lambda|^r). \tag{11}$$

It follows from (7) that

$$\min|\tilde{\lambda} - \lambda| \leqslant \max(r\theta, r^{1/r}\theta^{1/r}), \tag{12}$$

where θ is the right-hand side of (9) and $r^{1/r} < 1.45$; resembling a previous result [3]. Therefore, a multiple nonsemisimple eigenvalue λ will exhibit generally, under a perturbation δA of order ε, shifts proportional to $\varepsilon^{1/r}$; i.e. they are sensitive to changes in the matrix elements; being a known phenomenon both in numerical analysis and applied mathematics. For it is shown in [8], that if A is nonderogatory and has an eigenvalue λ of multiplicity m, then $\lambda(\varepsilon)$ of $A + \varepsilon A_1$ can be expanded into a Puiseux series of $\varepsilon^{1/m}$; namely

$$\tilde{\lambda}(\varepsilon) = \lambda + \sqrt[m]{\varepsilon}\lambda^{(1)} + \sqrt[m]{\varepsilon^2}\lambda^{(2)} + \cdots \tag{13}$$

in which $\lambda^{(1)}$ is obtained in [5] in the simple form

$$\lambda^{(1)} = \sqrt[m]{\langle y^1, \delta A x^1\rangle}, \tag{14}$$

where x^1 and y^1 are the only solutions to $(A - \lambda I)x = 0$ and $(A^{\mathrm{T}} - \lambda I)y = 0$, respectively. The case where λ has index r is similar to (14) except that it is given in terms of the rth root [6].

4. Local bounds

Both the Bauer–Fike result and the bound in (9) are valid for the whole spectrum. Thus a group of ill-conditioned eigenvalues affect the bound for the well-conditioned ones: being a serious drawback. This led Wilkinson to introduce his famous $1/s_i$ factors measuring the sensitivity of the ith eigenvalue alone, being equal to $1/|y^{i*}x^i|$ when both x^i and y^i are normalized. This follows from

$$\delta\lambda_i \approx y^{i*}\,\delta A\, x^i / y^{i*}x^i \tag{15}$$

representing approximately the shift in a distinct eigenvalue λ_i [6]. On the other hand if x^i and y^i are such that $\|x^i\|_2 = 1$ and $y^{i*}x^i = 1$, then $\|y^i\|_2$ becomes a measure of the sensitivity of λ [13]. If δA is big however, (15) cannot be taken as bounding the shift in λ_i. For this we provide the following result.

Theorem 4.1. *If λ_i is a simple eigenvalue of a nondefective matrix A with corresponding eigenvector x^i, the eigenvalue and eigenvector of $A + \delta A$ are given by $\lambda_i + \delta\lambda_i$ and $x^i + \delta x^i$, where*

$$|\delta\lambda_i| \leqslant |y^{i\mathrm{T}}|\,|\delta A| \sum_{j=1}^{n} |x^j| \tag{16}$$

and

$$|\delta x^i| \leqslant \sum_{j \neq i} |\alpha_j|\,|x^j| \tag{17}$$

where y^{i*} is the eigenrow of A corresponding to λ_i and

$$|\alpha| = (|\alpha_1|, \ldots, |\alpha_n|)^{\mathrm{T}} \leqslant (I - D_i|T^{-1}|\,|\delta A|\,|T|)^{-1} D_i |T^{-1}|\,|\delta A|\,|x^i| \tag{18}$$

where

$$D_i = \mathrm{diag}\left(\frac{1}{|\lambda_1 - \lambda_i| - |\delta\lambda_i|}, \ldots, O_i, \ldots, \frac{1}{|\lambda_n - \lambda_i| - |\delta\lambda_i|}\right) \tag{19}$$

provided that δA is such that

$$|\delta\lambda_i| + \|\,|T^{-1}|\,|\delta A|\,|T|\,\| \leqslant \min_{j \neq i} |\lambda_j - \lambda_i|. \tag{20}$$

Proof. From the eigenvalue problem

$$(A + \delta A)(x^i + \delta x^i) = (\lambda_i + \delta\lambda_i)(x^i + \delta x^i), \tag{21}$$

we get

$$A\delta x^i - \lambda_i \delta x^i = -\delta A x^i - \delta A\, \delta x^i + \delta\lambda_i x^i + \delta\lambda_i\, \delta x^i. \tag{22}$$

Let

$$\delta x^i = \sum_{j \neq i} \alpha_j x^j \tag{23}$$

then (22), for a particular $k \neq i$, and upon premultiplication by y^{k*}, reads

$$(\lambda_k - \lambda_i)\alpha_k = -y^{k*}\,\delta A x^i - \sum_{j \neq i} \alpha_j y^{k*}\,\delta A x^j + \delta\lambda_i\,\alpha_k \tag{24}$$

or that

$$|\alpha_k| \leqslant \frac{|y^{k\mathrm{T}}||\delta A||x^i| + \sum_{j \neq i}|y^{k\mathrm{T}}||\delta A||x^j||\alpha_j|}{|\lambda_k - \lambda_i| - |\delta\lambda_i|}, \quad k \neq i \tag{25}$$

also written as

$$(I - D_i|T^{-1}||\delta A||T|)|\alpha| \leqslant D_i|T^{-1}||\delta A||x^i|, \tag{26}$$

where D_i is given in (19). Thus

$$|\delta x^i| \leqslant |T||\alpha| \leqslant |T|(I - D_i|T^{-1}||\delta A||T|)^{-1}D_i|T^{-1}||\delta A||x^i| \tag{27}$$

if the spectral radius of the matrix $D_i|T^{-1}||\delta A||T|$ is less than one, or the more easily if

$$\min_{j \neq i}(|\lambda_j - \lambda_i| - |\delta\lambda_i|) \geqslant \||T^{-1}||\delta A||T|\| \tag{28}$$

as in (20). Note that (28) being satisfied implies from (9) that

$$\min_{j \neq i}|\lambda_j - \lambda_i| \geqslant |\delta\lambda_i| + \max_j|\delta\lambda_j| \tag{29}$$

or that the separation is larger than the sum of the perturbations; i.e. if the disks containing the perturbed eigenvalues are isolated.

As for the shift in the eigenvalue, we have from (7) that

$$|\tilde{\lambda} - \lambda_i||z_i| \leqslant \sum_{j=1}^{n}|y^{i\mathrm{T}}||\delta A||x^j||z_j|, \quad i = 1, \ldots, n \tag{30}$$

for any perturbed eigenvalue $\tilde{\lambda}$. Let $|z_k|$ be the maximum among $|z_i|$, then

$$|\tilde{\lambda} - \lambda_k||z_k| \leqslant \left(\sum_{j=1}^{n}|y^{k\mathrm{T}}||\delta A||x^j|\right)|z_k|, \tag{31}$$

i.e. that the union of the disks

$$Q_i = \left\{w\colon |w - \lambda_i| \leqslant \sum_j |y^{i\mathrm{T}}||\delta A||x^j|\right\}, \quad i = 1, \ldots, n \tag{32}$$

contains all characteristic roots of $A + \delta A$. So if the disks are isolated (32) bounds the shift $\delta\lambda_i$. □

5. Accuracy of computation

The bounds in (32) and (27) can be used to check the accuracy of a computed eigenvalue $\hat{\lambda}_i$ and its associated eigenvector $\hat{x}^i$. By setting $|\delta A| \leqslant \mathrm{eps}\,|A|$ (floating-point scheme), we obtain that

$$|\tilde{\lambda}_i - \lambda_i| \leqslant \mathrm{eps}\,|y^{i\mathrm{T}}||A|\sum_j|x^j| \tag{33}$$

and

$$|\hat{x}^i - x^i| \leqslant \text{eps}\,|T|\,(I - \text{eps}\,D_i|T^{-1}|\,|A|\,|T|)^{-1}D_i|T^{-1}|\,|A|\,|x^i|. \tag{34}$$

Example. Consider

$$A = \begin{bmatrix} 2 & 10^9 & -2\times 10^9 \\ -10^{-9} & 5 & -3 \\ 2\times 10^{-9} & -3 & 2 \end{bmatrix}, \quad \lambda = 2, 1, 6,$$

$$T = \begin{bmatrix} 3 & 2 & 1 \\ 2\times 10^{-9} & 2\times 10^{-9} & 2\times 10^{-9} \\ 10^{-9} & 2\times 10^{-9} & -10^{-9} \end{bmatrix},$$

$$T^{-1} = \begin{bmatrix} 0.75 & -0.5\times 10^9 & -0.25\times 10^9 \\ -0.5 & 0.5\times 10^9 & 0.5\times 10^9 \\ -0.25 & 0.5\times 10^9 & -0.25\times 10^9 \end{bmatrix},$$

$$|T^{-1}|\,|A|\,|T| = \begin{bmatrix} 19 & 20 & 14 \\ 20 & 21 & 15 \\ 14 & 15 & 11 \end{bmatrix}.$$

Thus

$$|\hat{\lambda}_1 - \lambda_1| \leqslant 53\,\text{eps}, \qquad |\hat{\lambda}_2 - \lambda_2| \leqslant 56\,\text{eps}, \qquad |\hat{\lambda}_3 - \lambda_3| \leqslant 40\,\text{eps}$$

or that the computed eigenvalues are expected to lose two significant digits at most. Such matrix has a well-conditioned eigenproblem; and yet we obtain very pessimistic results with either the Bauer–Fike bound: $|\hat{\lambda} - \lambda| \leqslant \text{eps}\,\|T^{-1}\|\,\|A\|\,\|T\| \approx \text{eps}\,10^{18}$ or the s_i factors of Wilkinson: $1/s_i \approx 10^9$. Obviously, the expression for $\delta\lambda_i$ in (15) — once its absolute value taken — corresponds to the element (i, i) in the matrix $|T^{-1}|\,|A|\,|T|$ of (9). But numerical analysts [13, 16] tend to use

$$|\delta\lambda_i| \leqslant \text{eps}\,\|y^i\|\,\|A\|\,\|x^i\|/|y^{i*}x^i|$$

as an indication of accuracy. Such expression can lead to erroneous results in pathological cases like it was shown above. Moreover, the above expression for $\delta\lambda_i$ as given from (15) is only valid for small perturbations, whereas (16) is valid for large δA.

For the shifts in the eigenvectors, we use (34), and for $|\hat{x}^1 - x^1|$ say,

$$D_1 = \operatorname{diag}(0, 1/1 - 53\,\text{eps}, 1/4 - 53\,\text{eps}),$$

$$x^1 = (1 \;\; 0.6666667\times 10^{-9} \;\; 0.3333333\times 10^{-9})^{\mathrm{T}}$$

and

$$|\hat{x}^1 - x^1| \leqslant \text{eps}(12 \quad 18 \times 10^{-9} \quad 12 \times 10^{-9})^{\mathrm{T}}.$$

We checked these bounds using the EGNRF subroutine of the IMSL library which confirmed the above results and which gave answers correct to six significant digits when executed on an eight-digit machine.

We note that (34) can also be written in the form

$$|\hat{x}^i - x^i| \leqslant \text{eps}(I - \text{eps}\,|T\,|D_i|\,T^{-1}|\,|A|)^{-1}|T\,|D_i|\,T^{-1}|\,|A|\,|x^i| \tag{35}$$

and in norm bound form

$$\frac{\|\hat{x}^i - x^i\|}{\|x^i\|} \leqslant \frac{\text{eps}\,\||T\,|D_i|\,T^{-1}|\,|A|\|}{1 - \text{eps}\,\||T\,|D_i|\,T^{-1}|\,|A|\|} \tag{36}$$

$$\leqslant \frac{\text{eps}\,\||T|\,|T^{-1}|\,|A|\|}{(\min_{j \neq i}|\lambda_j - \lambda_i| - |\delta\lambda_i|)\left(\dfrac{1 - \text{eps}\,\||T|\,|T^{-1}|\,|A|\|}{\min_{j \neq i}|\lambda_j - \lambda_i| - |\delta\lambda_i|}\right)}. \tag{37}$$

But since $|\delta\lambda_i| \leqslant \text{eps}\,\||T|\,|T^{-1}|\,|A|\|$ a result obtained like (10) from (6) namely $\det(I - T(\tilde{\lambda}I - \Lambda)^{-1}T^{-1}\,\delta A) = 0$, one finally gets that

$$\frac{\|\hat{x}^i - x^i\|}{\|x^i\|} \leqslant \frac{\text{eps}\,\||T|\,|T^{-1}|\,|A|\|}{\min_{j \neq i}|\lambda_j - \lambda_i| - 2\,\text{eps}\,\||T|\,|T^{-1}|\,|A|\|} \tag{38}$$

$$= \frac{\psi}{\text{sep}_i - 2\psi}, \tag{39}$$

where ψ bounds the shifts in the eigenvalues and sep_i is the ith eigenvalue separation from the nearest eigenvalue. So if the eigenvalues are well separated the eigenpair $(\hat{\lambda}_i, \hat{x}^i)$ has comparable accuracy. A matrix has therefore an ill-conditioned eigenproblem if

$$\text{sep} < 2\psi \tag{40}$$

or that the disks containing $\hat{\lambda}_i$, $i = 1, \ldots, n$ intersect due to poor separation.

The case where A has multiple eigenvalues is not much different if A is nondefective or λ itself semisimple. If λ is of multiplicity m together with eigenvectors $x^1, \ldots, x^m$, then

$$u^i = \sum_{j=1}^{m} c_{ij}x^j, \quad i = 1, \ldots, m \tag{41}$$

is also an eigenvector. Whereas

$$\delta u^i = \sum_{j \neq 1, \ldots, m} \alpha_{ij}x^j \tag{42}$$

is the shift in the eigenvector u^i. Eqs. (33) and (34) become

$$|\delta\lambda| = |\hat{\lambda} - \lambda| \leqslant \text{eps} \max_{i=1,\ldots,m} |y^{i^T}||A| \sum_{j=1}^{n} |x^j| \tag{43}$$

and

$$|\hat{u}^i - u^i| \leqslant \text{eps}|T|(1 - \text{eps}\, D|T^{-1}||A||T|)^{-1} D|T^{-1}||A| \sum_{j=1}^{n} |c_{ij}||x^j|, \tag{44}$$

where $|c_{ij}| \leqslant 1$ upon normalizing u^i (assuming A real symmetric without loss of generality) and where

$$D = \text{diag}\left(0, \ldots, 0, \frac{1}{|\lambda_{m+1} - \lambda| - |\delta\lambda|}, \ldots, \frac{1}{|\lambda_n - \lambda| - |\delta\lambda|}\right). \tag{45}$$

The bound in (44) is valid for all $i = 1, \ldots, m$. Generally speaking it depends on c_{ij} existing with the last term of (44), and is obtained from solving the eigenvalue problem [6]

$$0 = -\sum_{j=1}^{m} c_{ij} y^{k^*} \delta A x^j - y^{k^*} \delta A \sum_{j \neq 1,\ldots,m} \alpha_{ij} x^j + \delta\lambda\, c_{ik}, \quad k = 1, \ldots, m. \tag{46}$$

For first-order perturbations, the second term is missing, and $(\delta\lambda I - \underset{\sim}{T}^{-1} \delta A \underset{\sim}{T})c = 0$ yields both $\delta\lambda$ and c where $\underset{\sim}{T}$ and $\underset{\sim}{T}^{-1}$ are $n \times m$ and $m \times n$ matrices of x^j and y^j, $j = 1, \ldots, m$. But since the second term exists for large perturbations, (46) cannot be used to obtain c_{ij}. However, as we said, $|c_{ij}| \leqslant 1$ from $\|u^i\|_E = \|x^i\|_E = 1$ and $\langle x^i, x^j \rangle = 0$ for real symmetric A. If A is not real symmetric, then from $c^* \underset{\sim}{T}^* \underset{\sim}{T} c = 1$, $|c_{ij}| \leqslant 1/\sigma_n(\underset{\sim}{T})$ where $\sigma_n(\underset{\sim}{T})$ is the smallest singular value of $\underset{\sim}{T}$.

6. Bounds for matrix polynomials

For the eigenvalue problem

$$L(\lambda)x = (A_n\lambda^n + A_{n-1}\lambda^{n-1} + \cdots + A_0)x = 0, \tag{47}$$

we give a bound for the shift in the eigenvalue λ under perturbation δA_i, $i = 0, \ldots, n$. This problem was studied in [7] for small order perturbations $\delta A_i = \varepsilon \bar{A}_i$ in which the eigenvalue $\tilde{\lambda}(\varepsilon)$ of

$$L(\tilde{\lambda})\tilde{x} = [(A_n + \varepsilon\bar{A}_n)\tilde{\lambda}^n + (A_{n-1} + \varepsilon\bar{A}_{n-1})\tilde{\lambda}^{n-1} + \cdots + A_0 + \varepsilon\bar{A}_0]\tilde{x} = 0 \tag{48}$$

is obtained approximately as

$$\tilde{\lambda} \approx \lambda + \varepsilon\lambda^{(1)} \tag{49}$$

for λ simple or

$$\tilde{\lambda} \approx \lambda + \varepsilon^{1/m}\lambda^{(1)} \tag{50}$$

for λ multiple, and where $\lambda^{(1)}$ is exactly calculated in both cases. In this section, we provide a global bound similar to the one in Section 2.

Eq. (47) can be transformed into the generalized eigenvalue problem

$$\begin{bmatrix} 0 & I & & \\ & 0 & I & \\ & & \ddots & \\ -A_0 & -A_1 & \dots & -A_{n-1} \end{bmatrix} \begin{bmatrix} x \\ x^1 \\ \vdots \\ x^{n-1} \end{bmatrix} = \lambda \begin{bmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & A_n \end{bmatrix} \begin{bmatrix} x \\ x^1 \\ \vdots \\ x^{n-1} \end{bmatrix}, \tag{51}$$

i.e. briefly

$$Au = \lambda Bu \tag{52}$$

and we study the effect of δA and δB upon λ. We shall only consider the case where $\det(A_n) \neq 0$ ($L(\lambda)$ monic). Since

$$\det(\tilde{\lambda}[B + \delta B] - [A + \delta A]) = 0 \tag{53}$$

then

$$\det(I + (\tilde{\lambda} I - A)^{-1}(\tilde{\lambda}\,\delta B - \delta A)) = 0. \tag{54}$$

And with the spectral decomposition [6] of $(\tilde{\lambda} B - A)^{-1}$ in the form

$$(\tilde{\lambda} B - A)^{-1} = C_f(\tilde{\lambda} I - J)^{-1} D_f \tag{55}$$

in which C_f and D_f are eigenvectors and eigenrows of $(\lambda B - A)$ and where $(\tilde{\lambda} I - J)^{-1}$ possesses linear divisors only for simplicity (the case in which J has a Jordan block can be treated like in Section 3), it follows that

$$\det(I + C_f(\tilde{\lambda} I - J)^{-1} D_f(\tilde{\lambda}\,\delta B - \delta A)) = \det(I + (\tilde{\lambda} I - J)^{-1} D_f(\tilde{\lambda}\,\delta B - \delta A) C_f) = 0. \tag{56}$$

And following a similar approach like in Section 2, we get

$$|\tilde{\lambda} - \lambda| \leqslant \| |D_f| |\tilde{\lambda}\,\delta B - \delta A| |C_f| \|. \tag{57}$$

Further, by writing

$$\tilde{\lambda}\,\delta B - \delta A = (\tilde{\lambda} - \lambda)\,\delta B - \delta A + \lambda\,\delta B \tag{58}$$

it follows that

$$|\tilde{\lambda} - \lambda| \leqslant \frac{\| |D_f| |\lambda\,\delta B - \delta A| |C_f| \|}{1 - \| |D_f| |\delta B| |C_f| \|} \tag{59}$$

or finally in machine bound

$$|\tilde{\lambda} - \lambda| \leqslant \mathrm{eps} \frac{\| |D_f| (|\lambda| |B| + |A|) |C_f| \|}{1 - \mathrm{eps} \| |D_f| |B| |C_f| \|}. \tag{60}$$

It is preferable sometimes to obtain bounds in terms of the physical perturbations rather than transforming $L(\lambda)x = 0$ into a generalized eigenvalue problem. A classical example is the famous second-order mechanical system

$$(M\lambda^2 + R\lambda + K)x = 0, \tag{61}$$

which can be dealt with by first calculating the $n \times 2n$ and $2n \times n$ matrices C_f and D_f such that

$$(M\mu^2 + R\mu + K)^{-1} = C_f(\mu I - \Lambda)^{-1}D_f \tag{62}$$

if $\det(M) \neq 0$. Then from (56) and based upon a property of determinants [6, p. 20]

$$\begin{aligned} &\det(I + C_f(\tilde{\lambda}I - \Lambda)^{-1}D_f(\delta M\tilde{\lambda}^2 + \delta R\tilde{\lambda} + \delta K)) \\ &\quad = \det(I + (\tilde{\lambda}I - \Lambda)^{-1}D_f(\delta M\tilde{\lambda}^2 + \delta R\tilde{\lambda} + \delta K)C_f) = 0 \end{aligned} \tag{63}$$

and by expanding the function $\delta M\tilde{\lambda}^2 + \delta R\tilde{\lambda} + \delta K$ around λ by Taylor's series, it follows upon setting

$$\begin{aligned} \psi &= \| |D_f| |\delta M\lambda^2 + \delta R\lambda + \delta K| |C_f| \|, \\ \phi &= \| |D_f| |2\delta M\lambda + \delta R| |C_f| \|, \\ \theta &= \| |D_f| |\delta M| |C_f| \|, \end{aligned} \tag{64}$$

that

$$|\tilde{\lambda} - \lambda| \leqslant \frac{1-\phi}{2\theta}\left(1 - \sqrt{1 - \frac{4\theta\psi}{(1-\phi)^2}}\right). \tag{65}$$

It is interesting to note that when $\theta \to 0$, the bound in (59) is recovered.

Acknowledgement

The author wishes to thank the referee for his valuable suggestions.

References

[1] F. Bauer and C. Fike, Norms and exclusion theorems, *Numer. Math.* **2** (1960) 137–141.
[2] H. Baumgärtel, *Endlichdimensionale Analytische Störungstheorie* (Akademie Verlag, Berlin, 1972).
[3] K. Chu, Generalization of the Bauer–Fike theorem, *Numer. Math.* **49** (1986) 685–691.
[4] A. Deif, Sensitivity analysis from the state equations by perturbation techniques, *Appl. Math. Modelling* **5** (1981) 405–408.
[5] A. Deif, Nondiagonable systems under small perturbations, *Z. Angew. Math. Phys.* **33** (1982) 282–288.
[6] A. Deif, *Advanced Matrix Theory for Scientists and Engineers* (Gordon and Breach, London, 2nd ed., 1991).
[7] A. Deif and P. Hagedorn, Matrix polynomials subjected to small perturbations, *Z. Angew. Math. Mech.* **66** (1986) 403–412.
[8] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1976).
[9] U. Kulisch and W. Miranker, The arithmetic of the digital computer. A new approach, *SIAM Rev.* **28** (1986) 1–40.
[10] P. Lancaster, *The Theory of Matrices* (Prentice-Hall, Englewood Cliffs, NJ, 1966).
[11] M. Pease, *Methods of Matrix Algebra* (Academic Press, New York, 1965).
[12] H. Rosenbrock, Sensitivity of an eigenvalue to changes in the matrix, *Electron. Lett.* **1** (1979) 278–279.
[13] G. Stewart, *Introduction to Matrix Computations* (Academic Press, New York, 1973).
[14] G. Stewart and J. Sun, *Matrix Perturbation Theory* (Academic Press, New York, 1990).
[15] J. Tomovic, *Sensitivity Theory* (Elsevier, New York, 1969).
[16] J. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).